MOLECULAR AND CELLULAR BIOLOGY, Feb. 1990, p. 760–769
0270-7306/90/020760-10$02.00/0

Vol. 10, No. 2

Highly Conserved Upstream Regions of the α_1-Antitrypsin Gene in Two Mouse Species Govern Liver-Specific Expression by Different Mechanisms

JEAN J. LATIMER,[1]† FRANKLIN G. BERGER,[2] AND HEINZ BAUMANN[1]*

Department of Molecular and Cellular Biology, Roswell Park Memorial Institute, Buffalo, New York 14263,[1] and Department of Biological Sciences, University of South Carolina, Columbia, South Carolina 29208[2]

Received 26 June 1989/Accepted 2 November 1989

α_1-Antitrypsin (AT), the major elastase inhibitor in mammalian serum, is produced primarily in the liver. We have characterized AT gene structure and expression in the mouse species *Mus caroli*, which expresses high levels of AT in the kidneys as well as in the liver. Analysis of cDNA and genomic clones showed that the AT gene in *M. caroli* exhibits high sequence homology (>90%) to the gene in laboratory mice (*M. domesticus*) throughout the coding and 5′-flanking regions. Despite this extensive sequence conservation, the functional organization of *cis*-acting regulatory elements governing liver-specific expression is strikingly different between these species. Transient-transfection assays showed that the proximal region of the *M. caroli* promoter (i.e., between −120 and −2 relative to the transcriptional start site) is 10-fold more active than the analogous region of *M. domesticus* in driving the expression of an indicator gene in cultured liver cells. The increased activity of the proximal region of the *M. caroli* AT promoter appears to be the result of one or both of the two base substitutions at positions −46 and −48. The weak proximal promoter in *M. domesticus* is compensated for by the presence of upstream, liver-specific enhancers between −199 and −520; the analogous region in *M. caroli* is inactive. Thus, during the course of evolution, the modest 7% sequence divergence that has occurred between the 5′-flanking regions of the AT genes in these two species has generated distinct, yet equally effective, modes of hepatocyte-specific expression.

The molecular basis for the liver-specific expression of α_1-antitrypsin (AT), a major serine protease inhibitor that functions in the control of neutrophil elastase activity (4, 29), has been intensely investigated in humans and mice. For the human gene, maximal liver-specific expression requires 261 nucleotides of the 5′-flanking region, which can be subdivided into a distal element located between −261 and −210 and two proximal elements located between −137 and −37 (6, 8, 26). Although the proximal elements function only in hepatic cells, the distal element acts as a nonspecific enhancer (8). In striking contrast, the mouse gene, which is structurally similar to the human gene, requires 500 bases of the upstream region for full activity (12); several domains that function specifically in hepatic cells and that interact with hepatocyte-enriched DNA binding proteins have been identified within this region (7, 12). Thus, liver-specific expression of the AT gene is maintained by different mechanisms in the two species, reflecting the accumulation of regulatory mutations since the two species diverged.

With one known exception, all mammals express AT predominantly in the liver. The wild-derived mouse species *Mus caroli*, in sharp contrast to laboratory mice (*M. domesticus*), transcribes the AT gene in the kidneys as well as the liver (2, 19, 24). Renal AT expression in *M. caroli* is specific to tubule cells, where it is regulated by androgens during development (19). This species difference in tissue specificity is due to a *cis*-acting genetic element (2). The striking features of AT expression in *M. caroli* make it a useful model for studying the evolution of tissue-specific gene expression.

In the present paper, we describe more detailed analysis of the structure and expression of the AT gene in *M. caroli* and compare it with that in *M. domesticus* (17). Of major interest is the finding that although the 5′-flanking regions of the genes in the two species are quite similar in sequence, they behave very differently with regard to their functions in governing liver-specific AT expression. Thus, in the relatively short time (i.e., 4 million to 6 million years) since the separation of *M. caroli* and *M. domesticus*, major changes have occurred in the molecular mechanisms regulating AT gene expression.

MATERIALS AND METHODS

Animals. Inbred strains of *M. domesticus* (C57BL/6J, BALB/cJ, and DBA/2J) and random-bred *M. caroli* were maintained and provided by V. Chapman, Roswell Park Memorial Institute. All mice were used at 8 to 12 weeks of age.

Extraction and analysis of RNA. Total tissue RNA was isolated as described by Labarca and Paigen (18). In vitro mRNA-directed protein synthesis was performed in a reconstituted cell-free translation system (15) provided by W. Held, Roswell Park Memorial Institute. AT was identified among the nascent ^{35}S-labeled polypeptides by two-dimensional polyacrylamide gel electrophoresis (22) followed by fluorography (3).

Primer extension analysis of transcriptional start sites was performed by using the methods of Field et al. (9), with minor modifications. A synthetic oligonucleotide corresponding to the anti sense strand of the AT gene between nucleotides +48 and +68 was 5′ end labeled with [γ^{32}P]ATP in the presence of polynucleotide kinase. A 25-μl solution containing 1.5 ng of oligonucleotide, 10 μg of total RNA, 0.5 M piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 6.5), 3 M sodium chloride, and 25 mM disodium EDTA was

* Corresponding author.
† Present address: Laboratory of Radiobiology and Environmental Health, University of California, San Francisco, CA 94143.

heated at 90°C for 2.5 min and incubated at 45°C overnight. After ethanol precipitation, the nucleic acids were dissolved in 50 μl of 0.05 M Tris hydrochloride (pH 8.3) containing 2 mM dithiothreitol, 5 mM $MgCl_2$, and 40 mM KCl; avian myeloblastosis virus reverse transcriptase was added along with the four deoxynucleoside triphosphates (200 μM each). After 2 h at 42°C, RNA was removed by base hydrolysis, and the extension products were fractionated on a 20% acrylamide–7 M urea gel and visualized by autoradiography.

Extraction and analysis of DNA. Total tissue DNA was purified by the methods of Chapman et al. (5). For Southern blot analysis (28), 10 μg of DNA was digested with the appropriate restriction endonuclease as specified by the supplier. DNA fragments were fractionated on 0.8% agarose gels, transferred to nitrocellulose, hybridized to the appropriate nick-translated probe, and visualized by autoradiography. The blots were hybridized and washed at 65°C in 0.3 M NaCl.

Isolation of AT cDNA and genomic clones. Poly(A)$^+$ RNA from *M. caroli* liver was used to generate a cDNA library in λgt10 by the methods of Gubler and Hoffman (13). The library was screened with the AT-specific probe p1796 (1) or pJL2.5 (24). Two non-overlapping cDNA clones corresponding to the AT mRNA were isolated (see below). A *M. caroli* genomic library (provided by M. Edgell, University of North Carolina) was screened with p1796 DNA. Two overlapping clones corresponding to the *M. caroli* AT gene were isolated (see Fig. 3). DNA sequencing was performed by the dideoxy-chain termination method (25) with Sequenase (U.S. Biochemical Corp.).

Analysis of AT promoter function. Restriction enzyme fragments with the endpoints indicated in Fig. 6 and 7 were cloned into the *Hin*dIII *Nde*I sites of pSV0CAT, which contains a promoterless chloramphenicol acetyltransferase (CAT) gene. Transient transfections were performed with the human hepatoma cell line Hep3B (provided by G. Darlington) or HeLa cells (provided by L. Maquat). Cells were maintained in modified Eagle medium containing 10% fetal calf serum. The CAT construct to be tested (15 μg) was transfected by the calcium phosphate precipitation method (11). After 6 to 8 h cells were shocked with 20% glycerol and then incubated for an additional 36 h; extracts were prepared, and CAT activities were determined (10). Transfection efficiencies were monitored by cotransfection with 5 μg of plasmid pCH110, which contains and expresses the *lacZ* gene from *Escherichia coli* (14); CAT activities were normalized to the β-galactosidase levels.

RESULTS

Sequence of the AT mRNA from *M. caroli*. To determine the structure of *M. caroli* AT, we analyzed two clones from an *M. caroli* liver cDNA library. The clones, which corresponded to nonoverlapping 600- and 700-nucleotide cDNA segments, were sequenced and found to comprise a continuous 1,350-nucleotide stretch containing an open reading frame of 412 amino acids (Fig. 1A). The mRNA sequence is 87% similar to that from *M. domesticus* and 70% similar to that from humans. The predicted amino acid sequence shares 87% identity with the *M. domesticus* sequence and 62% identity with the human sequence. *M. caroli* contains an in-frame deletion of 15 bases near the 5′ end of the open reading frame, resulting in a deletion of 5 amino acids. Interestingly, five potential glycosylation sites exist within the *M. caroli* polypeptide (Fig. 1B), whereas only three are found within the human sequence (29). The reactive center, i.e., the conserved region that determines the protease specificity of the inhibitor (4, 29), differs by two amino acids between *M. caroli* and *M. domesticus* (Fig. 1B). One of these differences occurs at the P_1 position, which is the site of cleavage and is considered to be a primary determinant of specificity (4, 29). The *M. caroli* polypeptide, like the human polypeptide, contains a methionine residue at this position (Fig. 1B, residue 376), whereas the *M. domesticus* polypeptide contains a tyrosine residue (16). The functional significance, if any, of these changes is unknown. The *M. domesticus* and human proteins inhibit neutrophil elastase (4, 29), whereas the target for the *M. caroli* protein has not been determined.

Structure of the *M. caroli* AT gene. Earlier, we suggested that a single AT gene, expressed in both the liver and the kidneys, is present in *M. caroli* (19). This was based upon the relatively simple pattern of restriction endonuclease fragments for the AT gene, along with the identification of each fragment within a single AT gene clone. Further evidence consistent with this notion includes the finding of a restriction fragment length polymorphism segregating among *M. caroli* individuals. Some mice (denoted LL) contain a 5.6-kilobase *Eco*RI fragment, some (denoted SS) contain a 4.5-kilobase fragment, and some (denoted LS) contain both (Fig. 2A). Densitometric scanning indicated that each band in the LS individuals was half the intensity of that in either LL or LS individuals. The simplest interpretation of these results is that the 5.6- and 4.5-kilobase fragments represent alleles of a single structural gene. We have also found polymorphism in AT protein structure, as defined by the pI of the polypeptide synthesized in vitro (Fig. 2B). Both liver and kidney RNAs direct the synthesis of the same polymorphic form of AT, indicating that the same gene is active in both organs.

The *M. caroli* AT gene was isolated from a liver gene library as two overlapping clones spanning a combined distance of about 20 kilobases (Fig. 3). The locations of exons within the clones were determined by hybridization to cDNA fragments and by sequencing. Exon numbers, sizes, and positions are conserved between *M. caroli* and *M. domesticus* (17) (Fig. 3).

Primer extension was used to determine the location of the transcriptional start site for the AT gene in both the liver and the kidneys. A chemically synthesized oligonucleotide corresponding to nucleotides 48 to 68 within exon II (Fig. 1A) served as primer. Identical 68-base products were generated with *M. caroli* liver or kidney RNA, as well as with *M. domesticus* liver RNA (Fig. 4). No additional extension products could be detected, even after prolonged exposure of the autoradiogram. This contrasts with the situation in human liver, in which specific upstream start sites that correspond to transcripts produced from additional promoters are observed (23); these are due to monocytes that infiltrate the liver and that actively express the AT gene (23).

Comparison of AT promoter function in *M. domesticus* and *M. caroli*. The sequence of the 5′-flanking region, exon I, and the first 360 nucleotides of intron I of the *M. caroli* AT gene were determined and compared with the *M. domesticus* (12) and human (6) sequences (Fig. 5). The *M. domesticus* and *M. caroli* sequences are very similar in exon I and intron I, whereas the human sequence is quite different. Within intron I, the *M. caroli* gene contains a 46-base-pair insertion flanked by 4-base-pair direct repeats. All three species are quite similar up to position −126 within the 5′-flanking region. This conserved region contains a number of impor-

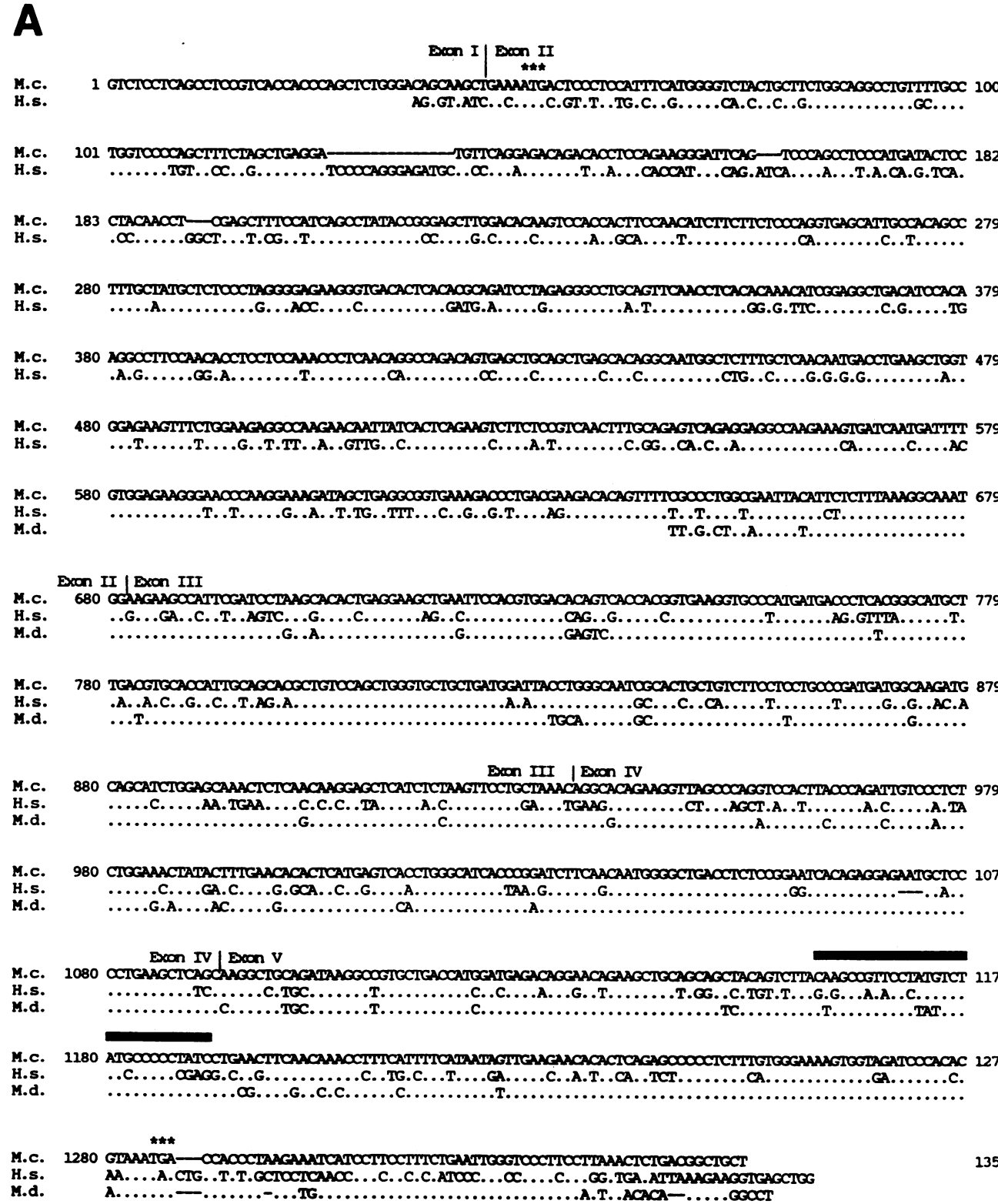

FIG. 1. Sequence of the *M. caroli* AT mRNA and polypeptide. (A) The sequence of the AT cDNA from *M. caroli* (M.c.) is shown aligned with those from humans (H.s.) (20) and from *M. domesticus* (M.d.) (16). The *M. caroli* sequence from nucleotides 1 to 20 (the cDNAs did not extend far enough) and positions of exon boundaries were derived from analysis of genomic clones (Fig. 5). Numbering begins at the transcriptional start site (Fig. 4). Symbols: •, identical nucleotides; −, deletions; *, initiation and termination codons for translations; ▬, region encoding the reactive center. (B) The amino acid sequences of the AT polypeptides, determined from cDNA sequences, are shown. Abbreviations and symbols are the same as in panel A. Potential glycosylation sites in *M. caroli* AT sequences are overlined.

B

```
M.c.   1 MIPSISWGLLLLAGLFCLVPSFLAEDVQE-----IDISRRDSVPASHD-  43
H.s.     .PS.V...I......C....VS....P.GDAAQK....HH.QDHPTFNKI

M.c.  44 TPYNLELSISLYRELGHKSNTSNIFFSQVSIATAFAMLSLGEKGDTHTQI  93
H.s.     ..NLA.FAF....Q.A.Q..ST.....P.............T.A...DE.

M.c.  94 LEGLQFNLTQTSEADIHKAFQHLLQTLNRPDSELQLSTGNGSLLNNDLKL 143
H.s.     ....N....EIP..Q..EG..E..H...Q...Q...T....LF.SEG...

M.c. 144 VEKFLEEAKNNYHSEVFSVNFAESEEAKKVINDFVEKGTQGKIAEAVKDP 193
H.s.     .D....DV.KL....A.T...DGT.....Q...Y.........VDL..EL

M.c. 194 DEDTVFALANYILFKGKWKKPFDPKHTEEAEFHVDTVTTVKVPMMTLTGM 243
H.s.     .R......V...F.....EF..EV.D...ED....Q.........KRL..
M.d.          SP..............EN.........ES..........S..

M.c. 244 LDVHHCSTLSSWVLLMDYLGNRTAVFLLPDDGKMQHLEQTLNKELISKFL 293
H.s.     FNIQ..KK........K....A..I.F...E..L....NE.THDI.T...
M.d.     ..................A..A...................S........

M.c. 294 LNRHRRLAQVHLPRLSLSGNYTLNTLMSHLGITRIFNNGADLSGITEENA 343
H.s.     E.EV..S.SL...K..IT.T.D.KSILGQ....KV.S.......V...-.
M.d.     ...R.....I.F....I..E.N.K....P.....................

M.c. 344 PLKLSKAADKAVLTMDETGTEAAAATVLQAVPMSMPPILNFNKPFIFIIV 393
H.s.     .......VH.....I..K.....G.MF.E.I...I..EVK.....V.LMI
M.d.     .....Q.VH.....I.........V...L...Y......R.DH..L...F

M.c. 394 EEHTQSPLFVGKVVDPTRK 412
H.s.     .QN.K....M....N..Q.
M.d.     .................H.
```

FIG. 1—*Continued.*

tant protein-binding sites, including those for hepatocyte nuclear factor 1 (HNF1) and C/EBP (7, 8, 12); in addition, the TATA box and CAAT box are located here. Upstream sequences up to −520 are well conserved between *M. domesticus* and *M. caroli* (Fig. 5). Within this segment from the *M. domesticus* gene, a number of binding sites for nuclear proteins have been identified and postulated to play important roles in the liver-specific expression of the AT gene (7, 12). Several of these sites, including those for hepatocyte nuclear factor 3 (HNF3) and C/EBP, show 1- to 3-base-pair differences between *M. domesticus* and *M. caroli* (Fig. 5).

In view of the high degree of sequence similarity exhibited by the AT genes in the two mouse species, we expected them to use the same *cis*-acting elements in the regulation of liver-specific expression. To test this notion, we compared analogous fragments of the *M. domesticus* and *M. caroli* promoter regions for their ability to drive the expression of the CAT gene in hepatic cells. The various fragments were subcloned upstream of the CAT gene in plasmid pSV0CAT,

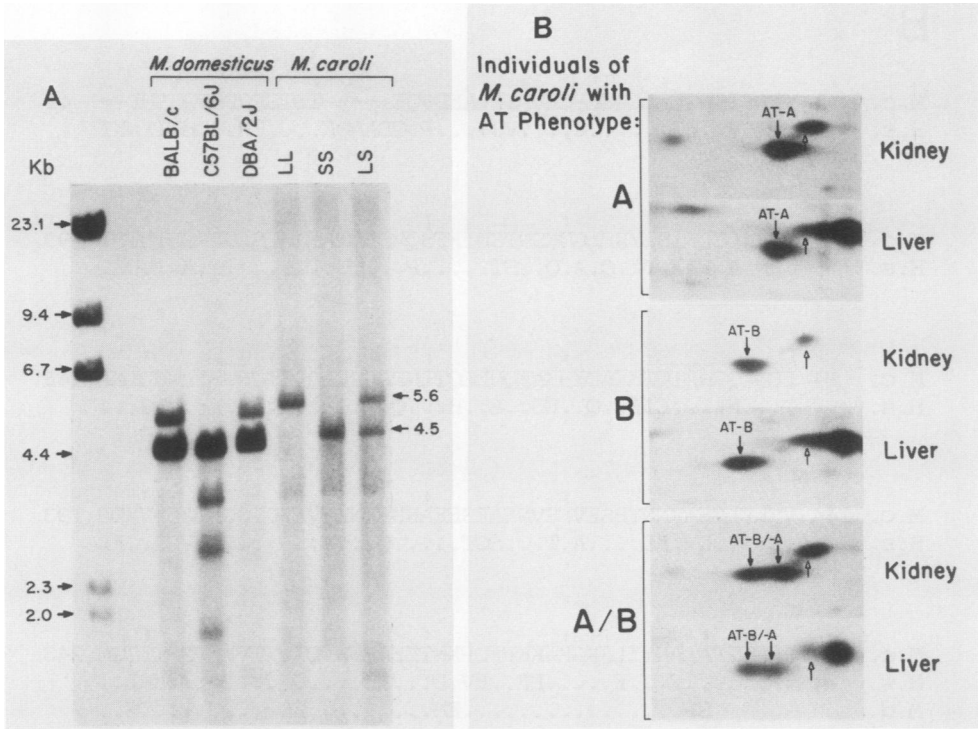

FIG. 2. Polymorphism in the AT gene from *M. caroli*. (A) Liver DNAs from *M. caroli* individuals and from several *M. domesticus* inbred strains were digested with *Eco*RI and analyzed by Southern blotting, with an *M. domesticus* exon I fragment (17) as the probe. (B) Total liver and kidney RNAs from three *M. caroli* individuals were translated in vitro, and the products were separated by two-dimensional gel electrophoresis. Symbols: ↓, AT forms A and B; ↑, actin polypeptide.

and resulting constructs were transfected into the human hepatoma cell line Hep3B; after 36 h, cell extracts were prepared and CAT activities were measured. Results were normalized to levels of β-galactosidase produced from co-transfected plasmid pCH110 (see Materials and Methods).

Figure 6 shows the results of transfection experiments with a variety of *M. domesticus* and *M. caroli* promoter fragments. *M. caroli* fragments −2700 to +86 and −520 to +86 (constructs 1 and 2, respectively) had levels of activity that were similar to *M. domesticus* fragment −522 to +47 (construct 6). Thus, maximal expression in each species requires about 520 base pairs of the 5′-flanking region. Deletion of the *M. domesticus* promoter from −522 to −199 resulted in 90% loss of activity (Fig. 6, compare constructs 6 and 7); this is similar to that reported by others (12) and indicates that the deleted region contains important elements governing AT promoter function. In striking contrast, deletion of the *M. caroli* promoter from −520 to −199 had no effect upon activity (Fig. 6, compare constructs 2 and 3). This differential effect of promoter deletions in the two species identifies a significant alteration in functional organization of the promoter regulatory elements. Further deletion to −140 reduced activity by about 50% in both species (Fig. 6, compare construct 3 with 4 and 7 with 8), indicating

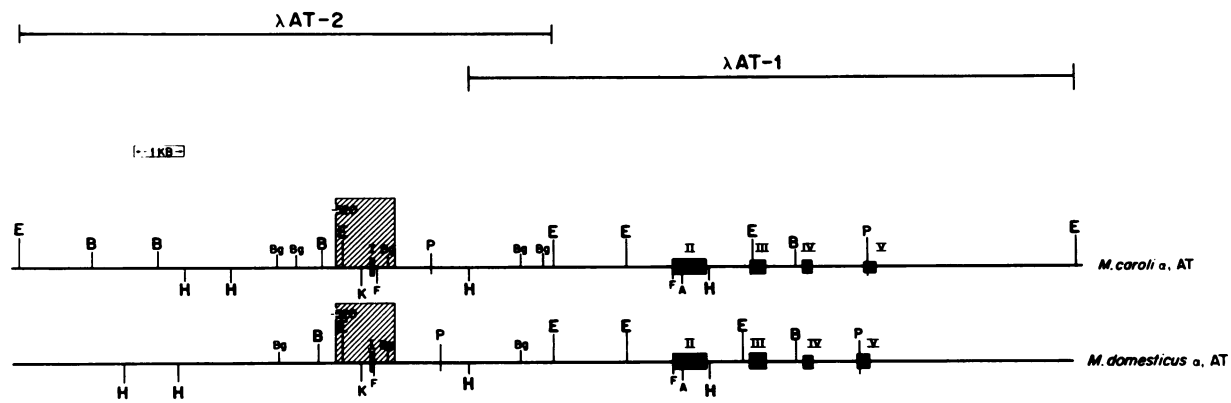

FIG. 3. Structure of the *M. caroli* AT gene. The restriction endonuclease map of the AT gene from *M. caroli* is shown aligned with that from *M. domesticus*, as determined previously (17). Symbols: ■, exons; ▨ (−520 to +370), promoter region. The promoter region was analyzed in detail (see Fig. 5 and text). Abbreviations: A, *Ava*II; B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; F, *Hin*fI; H, *Hin*dIII; P, *Pst*I.

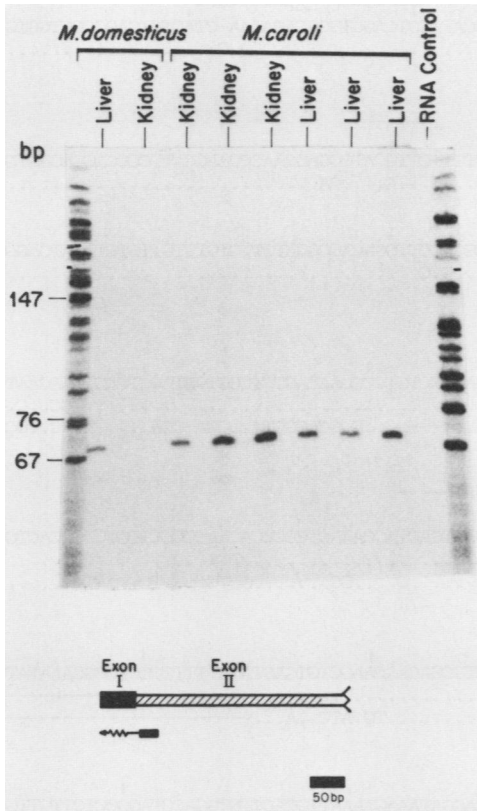

FIG. 4. Determination of the transcriptional start site for the *M. caroli* AT gene. An oligonucleotide corresponding to positions 48 to 68 of the AT mRNA was used in primer extension analysis of the *M. caroli* and *M. domesticus* liver and kidney mRNAs (see Materials and Methods for details). Extension products were analyzed by polyacrylamide gel electrophoresis. *Hpa*II-digested pBR322 DNA served as size markers. bp, Base pairs.

the presence of a positive regulatory element within this region. This may well represent effects of the previously identified HNF3-binding site between −199 and −181 (7). None of the constructs were active in HeLa cells (Fig. 7), indicating that the AT promoter in both species is liver specific.

A possible role for the 46-base-pair insertion within intron I of *M. caroli* was tested. Deletion of this region had no effect on promoter activity (Fig. 6, compare construct 4 with 5 or 11). In fact, the region between −140 and −2 maintained the 10-fold difference between species (Fig. 6, compare constructs 10 and 11), indicating that one or more of the seven single-base-pair differences within this segment are responsible for the species-specific alterations in activity of the proximal promoter.

Seven single-base-pair substitutions distinguish the *M. caroli* and *M. domesticus* AT genes between −140 and −2. These substitutions are clustered at two sites: five are between −133 and −122 and two are between −46 and −48 (Fig. 5). To determine which of the two clusters is responsible for the species differences in promoter activity, we tested additional 5′ deletions. Although deletion to −119 drastically reduced the activities for both AT constructs (Fig. 6, compare constructs 11 and 12 with 13 and 14), the activity of the *M. caroli* promoter was still several times higher than that of *M. domesticus*. We conclude from these results that AT promoter function in both species is significantly dependent upon sequences between −140 and −120 and that one or both of the two base substitutions at positions −46 and −48 are crucial for the elevated activity of the proximal region of the *M. caroli* promoter.

Thus, major functional differences distinguish the AT promoters in *M. domesticus* and *M. caroli*, despite the 93% sequence similarity. The region between −120 and −2 in *M. caroli* is 10-fold more active than that in *M. domesticus*; in addition, the upstream region between −520 and −199 in *M. domesticus*, but not in *M. caroli*, acts as a liver-specific enhancer of expression.

To test for functional independence of the differentially active distal and proximal segments of the AT promoters in *M. domesticus* and *M. caroli*, we examined chimeric constructs containing regions from both species. Construct 15 contains the distal region of *M. caroli* between −520 and −140 and the proximal region of *M. domesticus* between −140 and +47; conversely, construct 16 contains the distal region of *M. domesticus* between −522 and −140 and the proximal region of *M. caroli* between −140 and +86. The distal element of *M. caroli* stimulated the proximal region of *M. domesticus* about twofold (Fig. 6, compare constructs 8 and 15); this is most probably a consequence of the HNF3-binding site described above. However, the distal region of *M. domesticus* exerted nearly a 10-fold activation of the proximal region of *M. caroli* (Fig. 6, compare constructs 4 and 16). In fact, transfection with construct 16 gave a higher CAT activity than any of the other constructs; this is attributable to the positive effects of the *M. domesticus* upstream enhancers on the highly active proximal region of *M. caroli*.

Similar species-specific differences were observed when constructs were transfected into HepG2, a human hepatoma cell line. The results, taken together, indicate that in human hepatic cells the *M. caroli* and *M. domesticus* promoters, which are 93% homologous up to nucleotide −520, exhibit the same level of activity, yet specific regulatory elements within the promoters do not colocalize. The proximal region of the *M. caroli* promoter has a 10-fold-higher activity than the analogous region of *M. domesticus*, whereas the distal enhancer(s) of *M. domesticus* is not functional in *M. caroli*.

DISCUSSION

The unusual tissue specificity of AT expression in *M. caroli* (2, 19, 24) has prompted a detailed analysis of AT gene structure and expression in this species. In *M. caroli*, AT is produced from the same gene in both the liver and kidneys, using identical transcriptional start sites (Fig. 4). This differs from the situation in humans, where the extrahepatic expression (i.e., that in monocytes) is achieved through utilization of an upstream initiation site (23).

The amino acid sequence of *M. caroli* AT, deduced from cDNA sequencing, is 87% similar to that of *M. domesticus*. There are two differences within the reactive center. One occurs at the P_1 site, which is critical in defining inhibitor specificity (4, 29). Several AT genes within *M. domesticus* show 2% divergence, including a change in the P_1 amino acid residue (17). It is not clear whether the alterations that have been found within the reactive center of various AT polypeptides from *M. domesticus* and *M. caroli* affect the function of the inhibitors.

The high sequence homology between the *M. domesticus* and *M. caroli* AT genes, in addition to the fact that both species use the same transcriptional start site, would lead one to predict that the same *cis*-acting regulatory elements

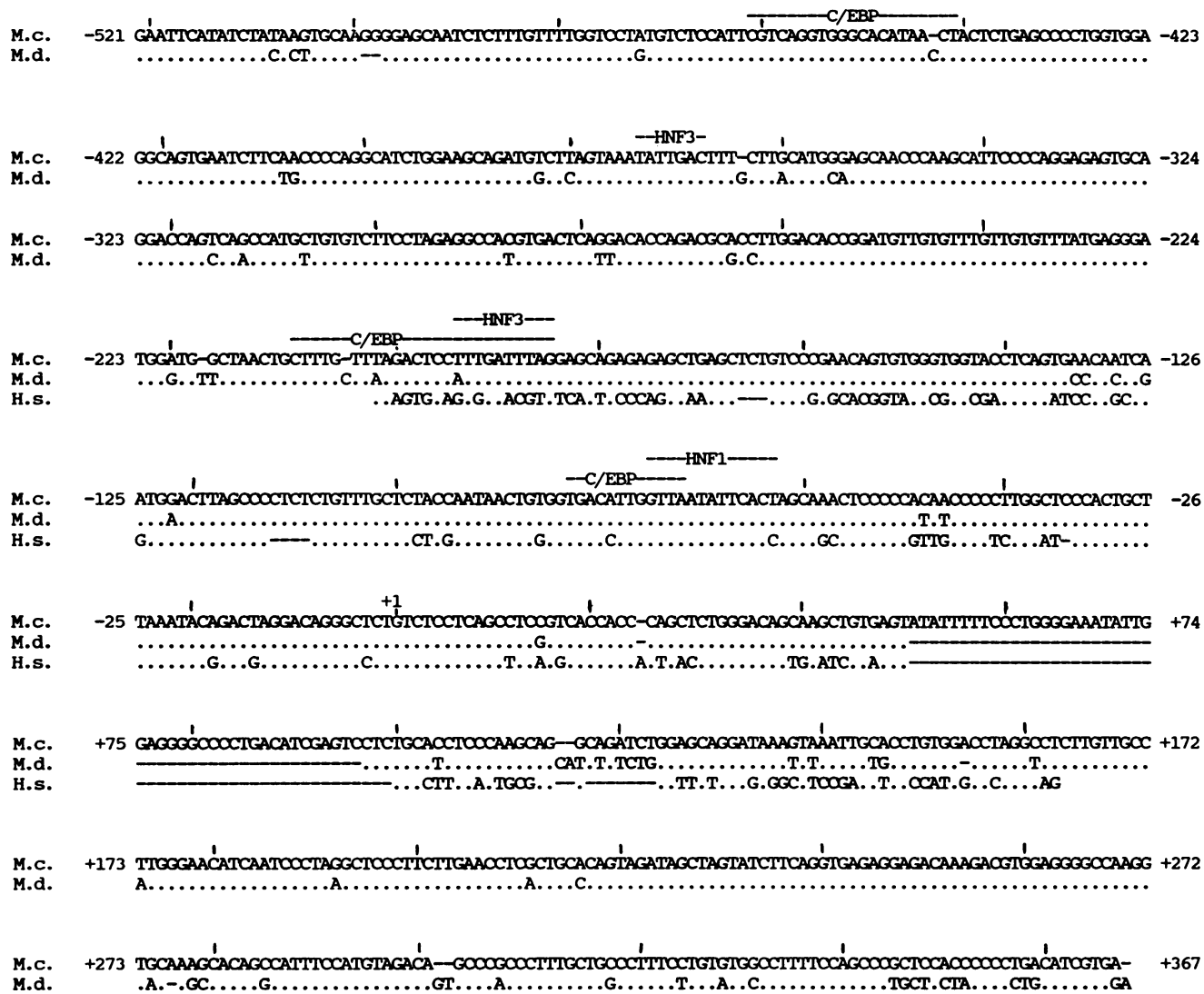

FIG. 5. Sequence of the promoter region of the *M. caroli* AT gene. The 5′-flanking region between nucleotides −521 and +367 of the *M. caroli* (M.c.) AT gene is aligned with the analogous region of the *M. domesticus* (M.d.) (12) and human (H.s.) (20) genes. The binding sites for C/EBP, HNF3, and HNF1 on the *M. domesticus* AT gene (7, 12) are indicated, along with the transcriptional initiation site (+1).

should be operating to control liver-specific expression. However, this prediction appears to be incorrect. Our present results clearly indicate that during the course of evolution, the modest 7% sequence divergence that has occurred between the 5′-flanking regions of *M. domesticus* and *M. caroli* has resulted in distinct, yet equally effective, modes of hepatocyte-specific regulation. Within the region between −120 and −2, only two base pairs distinguish the two species, yet promoter activities differ severalfold. The region between −199 and −520, which is 93% homologous between the species, functions as a positive enhancer in *M. domesticus*, yet is inactive in *M. caroli*. Thus, most of the regulatory information controlling *M. caroli* AT expression is located proximal to −199, whereas most of that controlling *M. domesticus* expression is located upstream of −199. Apparently, the presence of a weak proximal promoter in *M. domesticus* is compensated for by an upstream enhancer, so that the genes, in total, are transcribed with about equal efficiency in the two species.

It is now clear that several arrangements exist for the organization of regulatory elements determining the pattern of liver-specific AT expression. The human gene has been shown to be controlled primarily by the proximal region, along with modestly active nonspecific upstream enhancers (6, 8, 26). In contrast, the *M. domesticus* gene uses strong upstream liver-specific enhancers to activate the expression of a relatively weak proximal region (7, 12). Finally, as shown presently, the *M. caroli* gene utilizes the proximal region exclusively. The differences between *M. domesticus* and *M. caroli* are quite surprising, since the AT genes in the two species share extensive similarities in sequence. Regardless of molecular mechanism, the three strategies generate the same quantitative and developmental expression phenotypes in the liver.

The nature of the ancestral gene common to the various mouse AT genes is not known. It would be of interest to ascertain, for example, whether the *M. caroli* pattern or the *M. domesticus* pattern more closely represents the primordial gene. The fact that the *M. caroli* and human AT genes share a preference for the use of proximal regulatory ele-

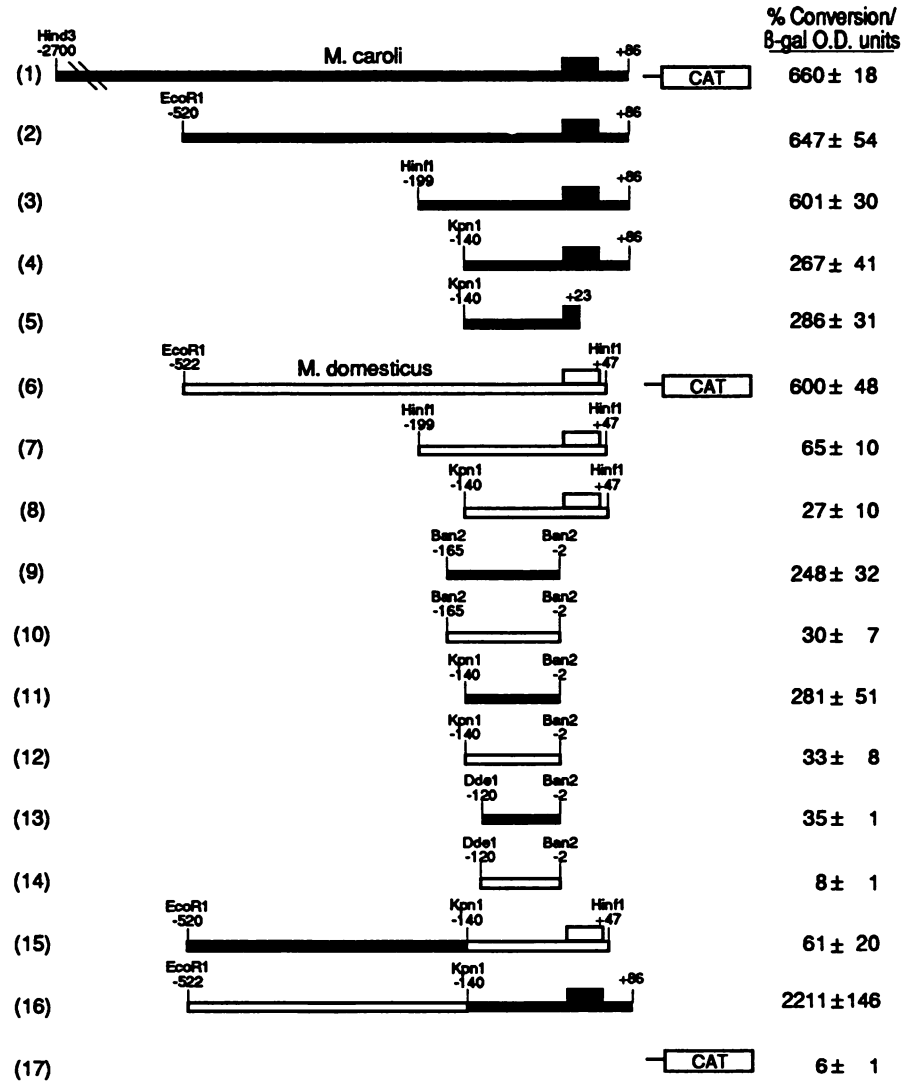

FIG. 6. Analysis of AT promoter function. The indicated constructs were transfected into Hep3B cells, and the specific CAT activities were determined. Values represent the average of three experiments (± standard deviation) and were normalized to β-galactosidase levels encoded by the cotransfected plasmid pCH110 (see Materials and Methods for details). O.D., Optical density.

ments suggests that this pattern may indeed by the ancestral type; thus, the recruitment of upstream enhancer function in *M. domesticus* may have evolved relatively recently, i.e., during the 4 million to 6 million years since its divergence from *M. caroli*.

The complex set of upstream elements governing AT expression in *M. domesticus* represent binding sites for regulatory factors such as HNF1, HNF3, and C/EBP (7, 12); the proximal region also contains sites for several of these factors (7). Some of the small number of sequence differences between 5'-flanking regions of the *M. domesticus* and *M. caroli* genes are within these protein-binding sites (Fig. 5) and therefore are candidates for determinants of the species difference. The C/EBP-binding site between nucleotides −464 and −445 contains a deletion of a C residue in the *M. caroli* gene; the HNF3 site between nucleotides −195 and −185, which functions as a hepatocyte-specific enhancer in HepG2 cells (7), contains an A-to-T change in *M. caroli*. Differences in regions outside currently known protein binding sites, such as those at −46 and −48, must also be important. These or perhaps other changes in the AT promoter influence protein binding and contribute to the species-specific differences in functional organization of the promoter.

The control of gene expression is increasingly being regarded as requiring multiple *cis*-acting elements that recognize a variety of different nuclear transcription factors (21). Thus, large nucleoprotein complexes are responsible for the expression phenotype of a particular gene. The AT gene, containing regulatory elements both proximal to and distal from the RNA start site, follows this emerging paradigm. Interestingly, the gene appears to contain multiple binding sites for the same transcription factor, suggesting that certain modular elements at some time may have undergone duplication and relocalization. The multiplicity of protein-DNA, as well as potential protein-protein, interactions that characterize the AT promoter provide a wide spectrum of targets for the evolutionary modification of promoter function. Some of these modifications alter the pattern of expression of the gene; others maintain the pattern.

Very little is known about the elements governing the

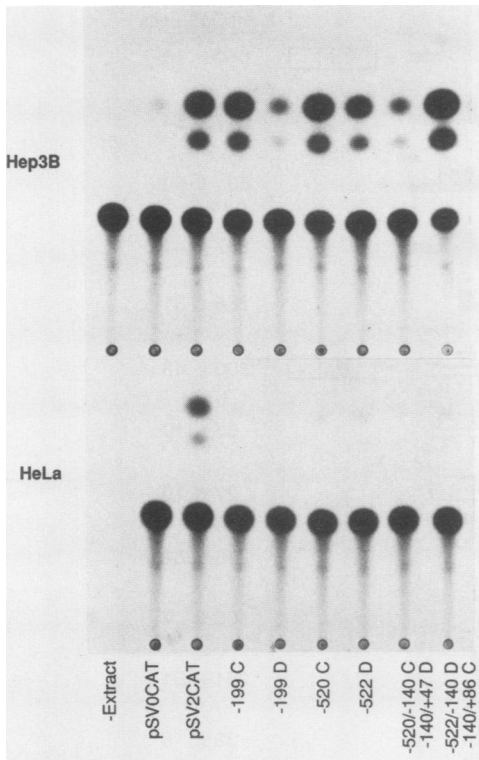

FIG. 7. Liver specificity of AT gene promoter function. Promoter constructs containing the −520 to +86 and −199 to +86 regions of the *M. caroli* AT gene (indicated by −520 C and −199 C) and the −522 to +47 and −199 to +47 regions of the *M. domesticus* gene (indicated by −522 D and −199 D) and reciprocal chimeric combinations of subregions were transfected into either Hep3B or HeLa cells. The CAT activities were determined by thin-layer chromatography.

uniquely high transcription rate of the AT gene in *M. caroli* kidney. Genetic studies indicate that these elements are *cis* acting (2). However, we have not obtained expression of *M. caroli* AT gene constructs in kidney cell lines (J. Latimer, unpublished data); thus, kidney-specific expression may be controlled by DNA sequences outside the region between −520 and +86. Interestingly, it has been observed that the human AT gene can be expressed at high levels in the kidneys of some, but not all, transgenic mouse lines (27). It is possible, therefore, that extrahepatic expression of the AT gene is controlled by negatively acting DNA elements; such elements may be absent, or their effects may be alleviated, in *M. caroli* and in some transgenic mouse strains. Clearly, more studies are necessary to clarify the nature of kidney-specific AT expression.

The physiological relevance of renal AT expression is unknown. Our present results, which show that low levels of sequence divergence within the regulatory region of the gene can have major effects on its function, raise the interesting possibility that kidney-specific expression is a passive consequence of selective forces driving alterations in DNA elements that regulate the liver phenotype; thus, once it occurs, renal expression may be too "costly" to remove and is therefore maintained without a physiological role.

ACKNOWLEDGMENTS

We are grateful to Karen Prowse for preparing several crucial promoter constructs. We thank Sanja Marinković, Gerald P. Jahreis, and Kwang-Ai Won for technical assistance and support; S. Grant for helpful discussion; and Mary Ketcham, Marcia Held, and Kristen French for secretarial assistance. We also thank K. Krauter and D. Grayson for providing genomic clones and B. Held for providing the in vitro translation system.

This work was supported by Public Health Service grant DK33886 from the National Institutes of Health. H.B. is an Established Investigator of the American Heart Association.

LITERATURE CITED

1. **Barth, R. K., K. W. Gross, L. C. Gremke, and N. D. Hastie.** 1982. Developmentally regulated mRNAs in mouse liver. Proc. Natl. Acad. Sci. USA **79**:500–504.
2. **Berger, F. G., and H. Baumann.** 1985. An evolutionary switch in tissue-specific gene expression: abundant expression of α_1-antitrypsin in the kidney of a wild mouse species. J. Biol. Chem. **260**:1160–1165.
3. **Bonner, W. M., and R. A. Laskey.** 1974. A film detection method for tritium-labeled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. **46**:83–88.
4. **Carrell, R., and J. Travis.** 1985. α_1-Antitrypsin and the serpins: variation and countervariation. Trends Biochem. Sci. **10**:20–24.
5. **Chapman, V., L. Forrester, J. Sanford, N. Hastie, and J. Rossant.** 1984. Cell lineage-specific undermethylation of mouse repetitive DNA. Nature (London) **307**:284–286.
6. **Ciliberto, G., L. Dente, and R. Cortese.** 1985. Cell-specific expression of a transfected human α_1-antitrypsin gene. Cell **41**:531–540.
7. **Costa, R. H., D. R. Grayson, and J. E. Darnell, Jr.** 1989. Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyetin and α_1-antitrypsin genes. Mol. Cell. Biol. **9**:1415–1425.
8. **DeSimone, V., G. Ciliberto, E. Hardon, G. Paonessa, F. Palla, L. Lundberg, and R. Cortese.** 1987. *cis*- and *trans*-acting elements responsible for the cell-specific expression of the human α_1-antitrypsin gene. EMBO J. **6**:2759–2766.
9. **Field, L. J., W. M. Philbrich, P. Howles, D. Dickinson, R. McGowan, and K. W. Gross.** 1984. Expression of tissue-specific Ren-1 and Ren-2 genes of mice: comparative analysis of 5′-proximal flanking regions. Mol. Cell. Biol. **4**:2321–2331.
10. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
11. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–461.
12. **Grayson, D. R., R. H. Costa, K. G. Xanthopoulos, and J. E. Darnell, Jr.** 1988. A cell-specific enhancer of the mouse α_1-antitrypsin gene has multiple functional regions and corresponding protein-binding sites. Mol. Cell. Biol. **8**:1055–1066.
13. **Gubler, U., and B. J. Hoffmann.** 1983. A simple and efficient method for generating cDNA libraries. Gene **25**:263–269.
14. **Hall, C. V., P. E. Jacob, G. M. Ringold, and F. Lee.** 1983. Expression and regulation of *Escherichia coli* lacZ gene fusions in mammalian cells. J. Mol. Appl. Genet. **2**:101–109.
15. **Held, W. A., K. West, and J. F. Gallagher.** 1977. Importance of initiation preparations in the translation of reovirus and globin mRNAs lacking a 5′-terminal 7-methylguanosine. J. Biol. Chem. **252**:8489–8497.
16. **Hill, R. E., P. H. Shaw, P. A. Boyd, H. Baumann, and N. D. Hastie.** 1984. Plasma protease inhibitors in mouse and man: divergence within the reactive center regions. Nature (London) **311**:175–177.
17. **Krauter, K. S., B. A. Citron, M.-T. Hsu, D. Powell, and J. E. Darnell, Jr.** 1986. Isolation and characterization of the α_1-antitrypsin gene of mice. DNA **5**:29–36.
18. **Labarca, C., and K. Paigen.** 1977. mRNA-directed synthesis of catalytically active mouse β-glucuronidase in Xenopus oocytes. Proc. Natl. Acad. Sci. USA **74**:4462–4465.
19. **Latimer, J. J., F. G. Berger, and H. Baumann.** 1987. Developmental expression, cellular localization, and testosterone regulation of α_1-antitrypsin in *Mus caroli* kidney. J. Biol. Chem. **262**:12641–12646.

20. **Long, G. L., T. Chandra, S. L. C. Woo, E. W. Davi, and K. Kurachi.** 1984. Complete sequence of the cDNA for human α_1-antitrypsin and the gene for the S variant. Biochemistry **23**:4828–4837.
21. **Maniatis, T., S. Goodbourne, and J. A. Fischer.** 1987. Regulation of inducible and tissue-specific gene expression. Science **236**: 1237–1244.
22. **O'Farrell, P. H.** 1975. High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. **250**:4007–4021.
23. **Perlino, E., R. Cortese, and G. Ciliberto.** 1987. The human α_1-antitrypsin gene is transcribed from two different promoters in macrophages and hepatocytes. EMBO J. **6**:2767–2771.
24. **Rheaume, C., J. J. Latimer, H. Baumann, and F. G. Berger.** 1988. Tissue- and species-specific regulation of murine α_1-antitrypsin gene transcription. J. Biol. Chem. **263**:15118–15121.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
26. **Shen, R. F., Y. Li, R. N. Sifers, H. Wang, C. Hardick, S. Y. Tsai, and S. L. C. Woo.** 1987. Tissue-specific expression of the human α_1-antitrypsin gene is controlled by multiple cis-regulatory elements. Nucleic Acids Res. **15**:8399–8415.
27. **Sifers, R. N., J. A. Carlson, S. M. Cliff, F. J. DeMayo, D. W. Bullock, and S. L. C. Woo.** 1987. Tissue-specific expression of the human alpha-1-antitrypsin gene in transgenic mice. Nucleic Acids Res. **15**:1459–1475.
28. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
29. **Travis, J., and G. S. Salvensen.** 1983. Human plasma protease inhibitors. Annu. Rev. Biochem. **52**:655–709.